Topological quantum computing with a noisy network and percent-level error rates

Naomi H. Nickerson[1], Ying Li[2], Simon C. Benjamin[2,3]

[1] *Department of Physics, Imperial College London, Prince Consort Road, London SW7 2AZ, UK*
[2] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543 and*
[3] *Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, UK.*
Please address correspondence to: *s.benjamin@qubit.org*

A quantum computer could be built by networking together many simple processor cells. The difficulty is that realistic quantum links are very error prone. A solution is for cells to repeatedly communicate with each other and so 'purify' any imperfections; however prior studies suggest that the cells themselves must then have extremely low internal error rates. Here we describe a method by which even error-prone cells can perform purification: groups of cells generate shared resource states, which then enable stabilization of topologically encoded data. Given a realistically noisy network ($\geq 10\%$ error rate) we find that intra-cell error rates for initialisation, state manipulation and measurement can simultaneously exceed 0.82% before the protocol fails.

Topological codes are an elegant and practical method for representing information in a quantum computer. The units of information, or logical qubits, are encoded as collective states in an array of physical qubits [1]. We employ stabilizers, i.e. measurements on nearby groups of physical qubits, to detect errors as they arise. By a suitable of choice stabilizers we can even generate logical operations between the encoded qubits. The act of applying stabilizers over the array thus constitutes a kind of 'pulse' for the computer – it is a fundamental repeating cycle and all higher functions are built upon it.

If the stabilizer operations themselves are perfect, then a topological code can be remarkably robust to noise; for example $\sim 19\%$ of physical qubits can suffer errors before the encoded information is corrupt [2]. However here we are concerned with the case that the stabilization process itself can suffer errors, through imperfections in steps such as initialising ancillas, performing gate operations and measurements. Stabilizers will still "do more good than harm", and extend the lifetime of encoded information, provided that such errors are sufficiently infrequent. Estimates of the threshold error rate are around $0.75\% - 1.4\%$ depending on the model [3–5].

These numbers are applicable to a monolithic architecture (Fig. 1 upper), i.e. a large grid of qubits with each connected to its neighbours such that two-qubit operations can be performed with high fidelity [5]. Building such a structure is a plausible goal for some systems, notably specific superconducting devices [6]. However it is less appropriate for other nascent quantum technologies. For example an individual nitrogen-vacancy (NV) centre in diamond, with its electron spin and small number of associated nuclear spins, naturally forms an isolated cluster [7]. Similarly a small ion trap holding a modest number of ions is a self-contained unit [8]. We use the term 'cell' for these convenient small groupings of qubits. Cells could be linked up to form a networked array (Fig. 1 lower), but we should expect links to have relatively poor fidelity. Links may be photonic, in which case a key issue

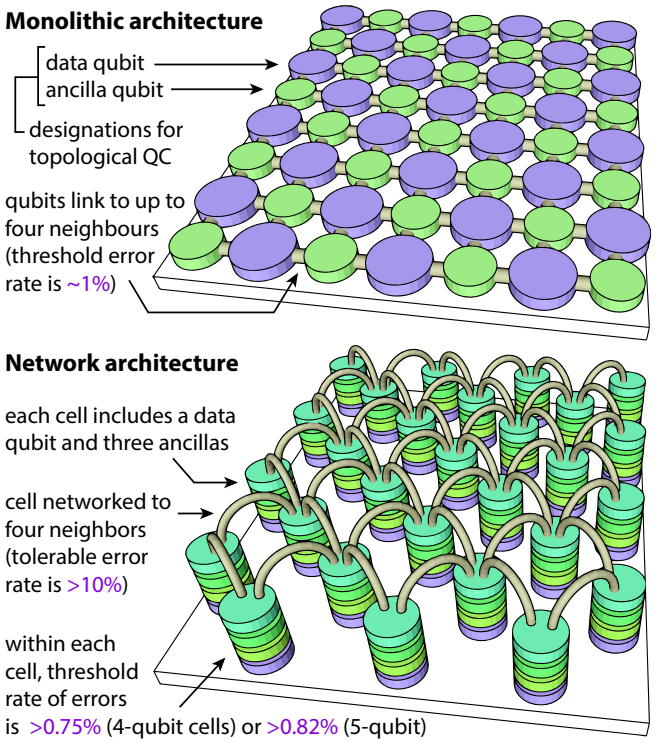

FIG. 1: Quantum architectures. Upper figure follows the scheme of Ref. [5]. These schematics are examples of very broad classes of systems; e.g. the network may actually be composed of NV centre 'cells' linked by spin chain 'wires'.

is photon loss, and instabilities in path lengths or interaction strengths can also degrade the channel. An alternative approach in the solid state is to employ a chain of spins to 'wire' together different cells [9]. Here again it may be infeasible to achieve error rates below the 1% level, given realistic decoherence affecting the spins [10].

Fortunately the problem of poor inter-cell links is not insurmountable: cells can interact repeatedly and *purify* the results to remove errors [11, 12]. Thus one can realise

a 'noisy network' (NN) paradigm, which is also called *distributed* [13, 14] or *modular* [15]. However prior analyses of this approach have indicated a serious drawback. The need to perform purification means a cell's own internal error rate must be *very* low; of the 0.1% order [11–15]. This contrasts poorly with the 1% general error threshold for the monolithic case. Given that a real machine should operate well below threshold, such a demand on the intra-cell precision may be prohibitive.

Here we describe a new NN protocol which achieves a threshold for the intra-cell operations which is far higher, and comparable to current estimates for tolerable noise in monolithic architectures. We introduce an optimised network protocol enabling our cells to collectively represent logical qubits according to the 2D toric code (one of the simplest and most robust topological surface codes [1], directly related to the approach used in cluster state architectures [3, 13–15, 17]). This code is stabilized by repeatedly measuring the parity of groups of four qubits in either the Z or the X basis. We implement such measurements by directly generating a GHZ state shared across ancilla qubits in four cells, before consuming the resource in a single step to jointly stabilize one "data qubit" in each cell. This procedure is in the spirit of the bandaid protocol [16] and contrasts with the standard approach of performing a sequence of two-qubit gates between the four data qubits and a fifth auxiliary qubit (which, in the network picture, would require its own cell). By directly generating the GHZ state over the network we remove the need for the auxiliary unit and more importantly we can considerably reduce the accumulation of errors.

The protocol specified in Fig. 2 requires a total of four qubits within each cell of the network. One qubit in each cell is designated the data qubit. The other three are ancillas that will be initialised, processed and measured during each stabilizer evaluation process. This proves sufficient to handle network error rates at the 10% level while still achieving a near-percent level threshold for intra-cell error rates. We note that if a further ancilla qubit is available in each cell then one can tolerate still higher levels of network noise, or alternatively one can boost the intra-cell threshold even closer to 1% (please see appended Sup. Mat. document).

We use the network channel between two cells to create shared, noisy Bell pairs in the Werner form

$$\rho_{\text{raw}} = \left(1 - \frac{4p_n}{3}\right)|\Phi^+\rangle\langle\Phi^+| + \frac{p_n}{3}\mathbb{1}$$

where $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. Throughout this paper we choose the network error probability to be $p_n = 0.1$. This 'raw' Bell generation is the sole operation that occurs over the network. We additionally require only three intra-cell operations: controlled-Z (i.e. the two-qubit phase gate), controlled-X (i.e. the control-NOT gate) and single qubit measurement in the X-basis. The two

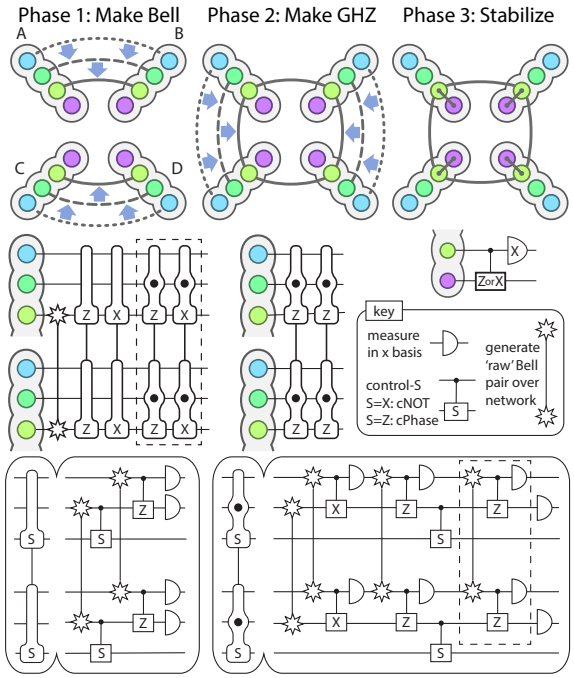

FIG. 2: The three phases for generating and consuming a 4-cell GHZ state among the ancilla qubits (blue-to-green) in order to perform a stabilizer measurement on the data qubits (purple). Phase (1): use purification to create a high quality Bell pair shared between cell A and cell B, while in parallel doing the same thing with cells C and D. (2) Fusion operations, A to C and B to D, create a high fidelity GHZ state. (3) Finally use the GHZ to perform a one-step stabilizer operation. The parity of the four measured classical bits is also the parity of the stabilizer operation we have performed on the data qubits. Two dashed regions indicate operations that are part of the STRINGENT protocol; omitting them yields the EXPEDIENT alternative.

qubit gates are modelled in the standard way: an ideal operation followed by the trace-preserving noise process

$$N(\rho) = (1 - p_g)\rho + \frac{p_g}{15}\sum_{A,B}(A \otimes B)\rho(A \otimes B)^\dagger$$

where operator $A \in \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ acts on the first qubit, and similarly B acts on the second qubit, but we exclude the case $\mathbb{1} \otimes \mathbb{1}$ from the sum. Meanwhile noisy measurement is modelled by perfect measurement preceded by inversion of the state with the probability p_m. Phases 1 and 2 in Fig. 2 are *postselective*: whenever we measure a qubit, one outcome indicates 'continue' and the other indicates that we must recreate the corresponding resource (the desired outcome is that which we would obtain, were the noise parameters set to zero). We note that in the circuits developed here we adopt and extend the "double selection" concept introduced by Fujii and Yamamoto in Ref. [18]. The Figure shows two variants of our protocol. EXPEDIENT minimises the number of raw pairs required and thus the overall time requirement. On average it requires fewer than twice the number of pairs consumed in

the ideal case that all measurements "succeed". Alternatively the STRINGENT variant imposes higher fidelity standards at the cost of a longer procedure. The performance of both protocols is found presently.

Having specified the stabilizer protocols, we must determine their real effect given the various error rates p_n, p_g and p_m. It is convenient derive a single superoperator describing the action of the entire stabilizer protocol on the four data qubits. In the following we write M to stand for the "odd" or "even" reported outcome of the stabilizer protocol (i.e. the parity of the four measured qubits in Phase 3 of Fig. 2). If density matrix ρ represents the state of all data qubits prior to the evaluation of our stabilizer, then measurement outcome M and the corresponding state $S^M(\rho) = P^M(\rho)/\mathrm{Tr}[P^M(\rho)]$ will occur with probability $\mathrm{Tr}[P^M(\rho)]$, for some projective operator $P^M(\cdot)$. Now stabilising the toric code involves two types of operation, the X and the $Z-$stabilizers. Suppose we are performing a Z stabilizer. If our protocol could act perfectly, then for example the *even* projector would be $P^{\mathrm{even}}_{\mathrm{ideal}}(\rho) = \sum |i\rangle\langle i|\rho|j\rangle\langle j|$, where the sum is over all states $|i\rangle$, $|j\rangle$ with definite even parity in the Z-basis: $|0000\rangle$, $|0011\rangle$, etc. Analogously the ideal *odd* Z-stabilizer sums over states of definite odd parity, and meanwhile for X stabilizers the ideal projectors refer to states of definite parity in the X-basis. In reality our imperfect operations result in projectors of the form

$$P^M_{\mathrm{real}}(\rho) = \sum_e a_e^M E_e P^M_{\mathrm{ideal}}(\rho) E_e^\dagger + b_e^{\bar{M}} E_e P^{\bar{M}}_{\mathrm{ideal}}(\rho) E_e^\dagger$$

with $E_e = (ABCD)_e$ and $\{A, B, C, D\} \in \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$.

Here the four operators making up E are understood to act on data qubits 1 to 4 respectively, and index e runs over all their combinations. The symbol $\bar{M}$ represents the compliment of M, i.e. "odd" for M ="even" and vice versa. We see that this real projector is made up of a mix of the correct and incorrect projectors together with possible Pauli errors; the various weights a and b capture their relative significance.

Given the underlying error rates p_n, p_g and p_m we can employ the Choi-Jamiolkowski isomorphism to find the corresponding weights $\{a, b\}$ in our superoperator. In the Supplementary Material we give examples. We find that the largest contributor to $P^M_{\mathrm{real}}(\rho)$ is the reported parity projection, i.e. for M ="even", the largest of all weights $\{a, b\}$ is that associated with $E = (\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1})$ and $P^{\mathrm{even}}_{\mathrm{ideal}}(\rho)$. The next largest term will be the pure "wrong" projection, i.e. the combination of $E = (\mathbb{1}\mathbb{1}\mathbb{1}\mathbb{1})$ and $P^{\mathrm{odd}}_{\mathrm{ideal}}(\rho)$. This form of error is relatively easy for the toric code to handle, and we have deliberately favoured it over other error types by minimising the σ_x and σ_y errors, rather than σ_z, in our stabilizer protocol. The remaining terms correspond to Pauli errors in combination with either the correct or the incorrect projectors; for example $E = (\sigma_x\mathbb{1}\mathbb{1}\mathbb{1})$ is an erroneous flip on data qubit 1. Such single-qubit errors are more probable than two-qubit errors, and the occurrence of three or four errors is highly improbable.

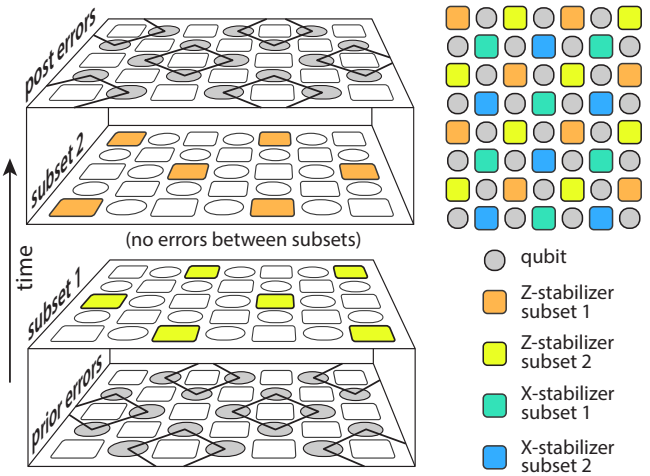

FIG. 3: Scheduling stabilizer operations. The right side graphic shows the standard arrangement of one complete stabilizer cycle, involving Z and X projectors (square symbols indicate that the four surrounding data qubits are to be stabilised). Because a given cell can only be involved in generating one GHZ resource at a given time, each of these two stabilizer types must be broken into two subsets; see main figure. Fortunately in our stabilizer superoperator $S(\rho)$ we can commute projectors and errors so as to expel errors from the intervening time between subsets, so allowing them to merge.

Having determined the stabilizer superoperators $P^M_{\mathrm{real}}(\cdot)$ for both Z-type and X-type stabilizers, together with a suitable scheduling scheme as shown in Fig. 3, we proceed to determine thresholds by intensive numerical modelling. Our model introduces errors randomly but with precisely the correlation rates indicated by the weights $\{a, b\}$. We record the (noisy) stabilizer outcomes and subsequently employ Edmonds' minimum weight perfect matching algorithm [19, 20] to pair and resolve stabilizer flips in the standard way (see e.g. Ref. [5]). This technique allows us to establish the threshold for successful protection of quantum information by simulating networks of various sizes. If an increase of the network size allows us to protect a unit of quantum information more successfully, then we are below threshold. Conversely if increasing the network size makes things worse, then in effect the stabilizers are introducing more noise than they remove and we are above threshold. The results of a large number of such simulations are shown in Fig. 5. We see thresholds for EXPEDIENT and for STRINGENT of over 0.6% and 0.77% respectively. The third graph shows the performance of a five-qubit-per-cell variation of EXPEDIENT which we describe in more detail in the Supplementary Material. It achieves a threshold in excess of 0.82% by employing an additional filter such that most stabilizers are improved while a known

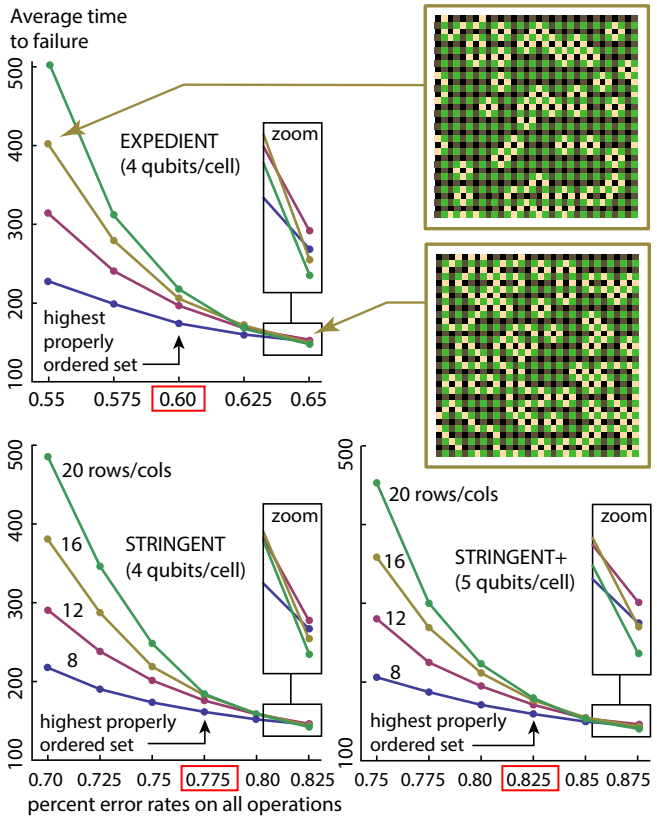

FIG. 4: Performance of the EXPEDIENT and STRINGENT network protocols, given a toric code with n rows $\times$ n columns ($2n^2$ qubits) of data qubits. A given numerical experiment is a simulation of 100 complete stabilizer cycles on an initially perfect array, after which we attempt to decode a Z measurement of the stored qubit. The result is either a success or a failure; for each data point we perform at least 10,000 experiments to determine the fail probability, and reciprocate this to infer an expected time to failure. Network error rates are 10% in all cases; we set intra-cell gate and measurement error rates equal, $p_m = p_g$, and plot this on the horizontal axes. For low error rates the system's performance improves with increasing array size. As the error rates pass the threshold this property fails. Insets: typical final states of the toroidal array after error correction. Yellow squares are flipped qubits, green squares indicate the pattern of Z-stabilizers. Closed loops are successful error corrections; while both arrays are therefore successfully corrected, it is visually apparent that the above-threshold case is liable to long paths.

minority become more noisy, and this classical information is fed into an enhanced Edmonds' algorithm.

The Edmonds' algorithm approach employed here has been chosen because it has been well studied in the context of noisy stabilizers; in the Supplementary Material we apply our same model to monolithic architecture, obtaining a threshold in the region of 0.9%, generally consistent with prior studies [5]. We note that for the idealised case of perfect (noiseless) stabilizers other algorithms have been developed which may offer advantages over Edmonds' [2, 21, 22]. Such approaches might eventually translate to higher thresholds for noisy stabilizers (both for monolithic and noisy network cases).

We have yet to consider general decoherence caused by the external environment during a stabilizer protocol. Fortunately many of the systems most relevant to the NN approach have excellent low-noise "memory qubits" available. For example in NV centres at room temperature, nuclear spins can retain coherence for the order of a second; and for impurities in silicon the record is several minutes [23, 24]. We would naturally use such spins for our data qubits and for the innermost ancilla qubits, i.e. those bearing the GHZ-state in Fig. 2. Then given that a conditional gate on a nuclear spin may take of order $10\mu s$, we estimate the probability of an environmentally induced error during a full stabilizer protocol to be at least an order of magnitude below the error rates due to p_g and p_m (see appended Sup. Mat. document).

In conclusion, we have described an approach to 'noisy network' quantum computing, where many simple modules or cells are connected with error prone links. We show that relatively high rates of error within each cell can be tolerated, thus closing the gap between error tolerance in networks versus monolithic architectures. We hope this result will encourage the several emerging technologies for which networks are the natural route to scalability; these include ion traps and NV centres (linked either optically or via spin chains).

We thank Ben Brown and Earl Campbell for numerous helpful conversations. This work was supported by the National Research Foundation and Ministry of Education, Singapore.

References

[1] E. Dennis, A. Kitaev, A. Landahl, J. Preskill, J. Math. Phys. **43**, 4452 (2002).
[2] H. Bombin *et al*, Phys. Rev. X **2**, 021004 (2012).
[3] R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007).
[4] D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Phys. Rev. A **83**, 020302(R) (2011).
[5] A. G. Fowler, M. Mariantoni, J. M. Martinis, A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).
[6] J. Ghosh, A. G. Fowler and M. R. Geller, arXiv:1210.5799.
[7] S. C. Benjamin, B. W. Lovett and J. M. Smith, Laser & Photon. Rev. **3**, 556 (2009).
[8] J. Benhelm, G. Kirchmair, C. F. Roos and R. Blatt, Nature Physics **4**, 463 (2008).
[9] N. Y. Yao *et al*, Nature Comm. **3**, 800 (2012).
[10] Y. Ping *et al*, arXiv:1210.6886.

[11] W. Dür and H.-J. Briegel, Phys. Rev. Lett. **90**, 067901 (2003).
[12] L. Jiang *et al*, Phys. Rev. A **76**, 062323 (2007).
[13] Y. Li and S. C. Benjamin, New J. Phys. **14** 093008 (2012).
[14] K. Fujii, *et al.* arXiv:1202.6588.
[15] C. Monroe *et al*, arXiv:1208.0391.
[16] K. Goyal, A. McCauley, and R. Raussendorf, Phys. Rev. A **74**, 032318 (2006).
[17] S. J. Devitt, A. M. Stephens, W. J. Munro and K. Nemoto, New J. Phys. **13**, 095001 (2011).
[18] K. Fujii and K. Yamamoto, Phys. Rev. A **80**, 042308 (2009).
[19] J. Edmonds, Canad. J. Math., **17** 449 (1965).
[20] Our code uses an implementation of the algorithm in V. Kolmogorov, Math. Prog. Comp. **1** 43 (2009) available from the author.
[21] G. Duclos-Cianci and D. Poulin, Phys. Rev. Lett. **104**, 050504 (2010).
[22] J. R. Wootton and D. Loss , Phys. Rev. Lett. **109**, 160503 (2012).
[23] P. C. Maurer *et al*, Science **336**, 1283 (2012).
[24] M. Steger *et al*, Science **336**, 1280 (2012).

Supplementary Material

COMPARISON TO THE MONOLITHIC CASE

In order to make a clear comparison with the threshold that can be achieved with the monolithic architecture, we applied our same superoperator description and numerical simulation to this case. (Of course there is no meaning for p_n, the network error rate, in a monolithic architecture). The figure shows the circuit we used for the case of a Z stabilizer. Note that this approach requires initialisation of the single shared auxiliary qubit; we used the value of p_m as the fidelity of this initialisation.

One sees that the threshold is between 0.9% and 0.95% and is therefore appreciably higher than that obtained by STRINGENT whilst in the same "ball park". As noted below, the addition of another ancilla per cell can further close this gap.

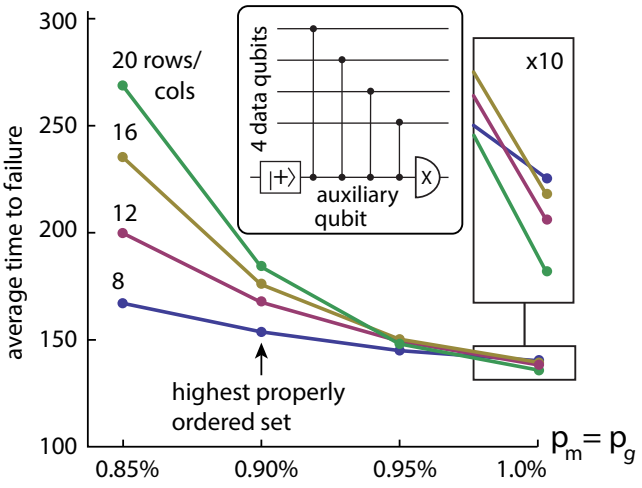

FIG. 5: Performance of the monolithic architecture, to be compared with the graphs in Fig. 4 of the main paper.

TWIRLING

The protocols given in the main paper lead to slight irregularities between the weights associated with given errors occurring on different specific data qubits; consequently in our superpoperator the weight associated with $ZZ\mathbb{1}\mathbb{1}$ might be slightly higher or lower than the weight associated with $\mathbb{1}ZZ\mathbb{1}$. This is because the protocol performs $A-B$, $C-D$ pairings in Phase 1, and then $A-C$, $B-D$ pairing in Phase 2, thus the errors that survive this purification process will not be equally distributed under rotation of cell labels. While these irregularities have no significance for our threshold calculations, they do cause the weightings to have a spurious complexity. Therefore in our analysis we append onto the procotols of

Fig. 2 a additional *twirling* operation which randomly applies swap operations between cells so as to 'smooth' the weightings. (Physically this is equivalent to the rather perverse act of programming one's system with a range of possible protocols, identical except for permutation of the cell labels A to D, and then applying one at random without retaining a record of the choice.)

EXAMPLE SUPEROPERATORS

In the main paper we wrote the following to represent the overall effect of the stabilizer protocol on four data qubits initially in state ρ,

$$P^M_{\rm real}(\rho) = \sum_e a_e^M E_e P^M_{\rm ideal}(\rho) E_e^\dagger + b_e^{\bar{M}} E_e P^{\bar{M}}_{\rm ideal}(\rho) E_e^\dagger \quad \text{with}$$
$$E_e = (ABCD)_e \text{ and } \{A,B,C,D\} \in \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}.$$

Here M stands for the reported outcome, "odd" or "even", and the $P^M_{\rm ideal}(\cdot)$ represents the perfect parity projector in either the Z or X basis depending on which class of stabilizer one is performing. The four operators making up E are understood to act on data qubits 1 to 4 respectively, and index e runs over all their combinations. The symbol $\bar{M}$ represents the compliment of M, i.e. "odd" for $M =$"even" and vice versa.

We noted that this real projector is made up of a mix of the 'correct' and 'incorrect' ideal projectors together with possible Pauli errors; the various weights a and b capture their relative significance. Here we give some examples of those weightings. The full list has 256 terms but they rapidly fall in magnitude so that the latter half are extremely small; here we list the terms corresponding to all single- and two-qubit errors. For the examples we will give, in fact the superscript M can be omitted from the a and b weights, i.e. the same set of weights a, b apply in $P^{\rm even}_{\rm real}$ as in $P^{\rm odd}_{\rm real}$. In other words the weight of the 'correct' projector $P^M_{\rm real}$ is the same in each, as are the weights of all the various error combinations. We will present the numbers as follows

$$\{\{A_{\mathbb{1}}, A_Z, A_X, A_Y, A_{ZZ}, A_{XX}, A_{YY}, A_{XZ}, A_{YZ}, A_{XY}\},$$
$$\{B_{\mathbb{1}}, B_Z, B_X, B_Y, B_{ZZ}, B_{XX}, B_{YY}, B_{XZ}, B_{YZ}, B_{XY}\}\}$$

Here the weights A are associated with the 'correct' projector $P^M_{\rm ideal}$ and the B weights are associated with the 'incorrect' projector $P^{\bar{M}}_{\rm ideal}$. The subscripts refer to the Pauli errors in the corresponding term, so that the subscript $\mathbb{1}$ indicates no errors, while subscript X means "a σ_X on *one of* the four data qubits", and YZ means "a σ_Y on one of the qubits and a σ_Z on another". These capital letters A and B are therefore sum of one or more of the weights a, b in Eqn. 1; for example $A_X = a_{\mathbb{1}\mathbb{1}\mathbb{1}X} + a_{\mathbb{1}\mathbb{1}X\mathbb{1}} + a_{\mathbb{1}X\mathbb{1}\mathbb{1}} + a_{X\mathbb{1}\mathbb{1}\mathbb{1}}$. As noted in the previous section, the twirling technique means that the individual a or b terms in such a sum in fact all have the same value.

For the case of EXPEDIENT operating with $p_n = 0.1$ and $p_m = p_g = 0.006$ we find projectors in the Z basis have following weights:

$$\{\{0.9117, 0.00681, 0.00314, 0.00314, 0.000182, 0.00000758,$$
$$0.00000758, 0.0000336, 0.0000336, 0.0000152\},$$
$$\{0.0617, 0.00674, 0.00314, 0.00314, 0.000127, 0.00000758,$$
$$0.00000758, 0.0000336, 0.0000336, 0.0000152\}\}.$$

(Meanwhile projectors in the X basis have the same numbers if we exchange subscripts $X \leftrightarrow Z$ in the listing order used above.) Note that many of the corresponding A and B terms are identical due to symmetries in the algorithm (including the twirling). We see that there is a 91% chance of a pure 'correct' projection, a 6.2% chance of a pure 'wrong' projection, a 0.68% chance of a correct projection followed by a single σ_Z error, and so on. For the case of STRINGENT operating with $p_n = 0.1$ and $p_m = p_g = 0.0075$ we find,

$$\{\{0.928, 0.00675, 0.00391, 0.00391, 0.000187, 0.00001182,$$
$$0.00001182, 0.0000414, 0.0000414, 0.0000236\},$$
$$\{0.0424, 0.00665, 0.00391, 0.00391, 0.0000849, 0.0000118,$$
$$0.00001182, 0.0000414, 0.0000414, 0.0000236\}\}.$$

Finally for the comparison case of the *monolithic* circuit where there is no network noise to consider, setting $p_m = p_g = 0.09$ we find

$$\{\{0.951, 0.00470, 0.00352, 0.00352, 0.00119, 0.0000128,$$
$$0.0000128, 0.00121, 0.00121, 0.0000256\},$$
$$\{0.0178, 0.00470, 0.00352, 0.00352, 0.00119, 0.0000128,$$
$$0.0000128, 0.00121, 0.00121, 0.0000256\}\}.$$

FIVE QUBITS PER CELL

The protocols in the main paper use a total of four qubits per cell of the network (one data qubit, three ancillas). If one adds further ancilla(s) then the performance improves. While it is beyond the scope of the present paper to seek optimal protocols that exploit five (or more) qubits, we can easily modify our approaches to make some use of the additional resource. For example, wherever our original 4-qubit STRINGENT protocol calls for a "raw" Bell pair to be created on a given ancilla pair we can instead insert a small circuit that creates raw pairs using the additional ancillas and purifies them. Similarly one can replace "single selection" purification (two tiers of ancilla yield an improved pair on the higher tier) with "double selection" (three tiers of ancillas yield

Level, L		Time steps, t_L	$p_{success,L}$	Failure reset level
1	Round one Bell pair production	7	0.7346	1
2	Round two Bell pair production	6	0.7506	1
3	Round one single selection	4	0.8619	3
4	Round two single selection	3	0.8550	3
5	Make GHZ	2	0.8651	1
6	Round one single selection	4	0.8619	6
7	Round two single selection	3	0.8550	7
8	Check GHZ	2	0.8654	1
9	Measure stabilizer	2	-	-

TABLE I: Listing of probabilities used in modelling the number of steps required for EXPEDIENT.

a significantly improved pair on the highest tier). Adopting these rather naive modifications we immediately find that the network noise is tolerable at the $p_n = 0.2$ level (rather than $p_n = 0.1$) for the same threshold of 0.77% local errors.

Alternatively we can hold the network errors constant and raise the intra-cell error rate. Data for this case are shown in the third panel of Fig. 4 in the main paper, labelled "STRINGENT+". We have introduced one further enhancement: We replace the usual Phase 3 of the STRINGENT protocol i.e. the steps where the GHZ resource would simply be coupled to the data qubits and then measured out. Instead we filter the GHZ *after* coupling it to the data qubits, and if this filter fails we *abort* the protocol by measuring the GHZ in the Z-basis (this therefore requires the addition of Z-basis measurement to our set of allowed primitive operations). If we have aborted then we then perform a whole new round of stabilization, this time without the filter so that the protocol will certainly complete. This procedure leaves us knowing the stabilizer outcome *and* some additional classical information about exactly what steps occurred. Specifically there are 3 distinct cases: (a) The filter was successfully passed; this is the best case and results in lower error rates on the data qubits – it is about 92% of cases in for the parameter range we consider. (b) The filter failed, but on measuring the GHZ in the Z-basis it was found to be in a correct state (e.g. 0000 or 1111). This occurs about 4% of the time. It is another "good" case in that there is a low chance of errors having reached the data qubits, so that the second round can perform normally. (c) The filter failed, and on measuring on the GHZ in the Z-basis it was found to be in an incorrect state (e.g. 0001). This occurs about 4% of the time and it is the "bad" case; in this event there is very likely to be an error on the data qubits. Now if we make no use of the classical information and merely "forget it" then the net effect of this protocol is to make things worse versus a simple one-round use of STRINGENT – this is not surprising since the overall risk of an error is not reduced (it is increased slightly due to the extra steps). However if we modify our Edmonds matching algorithm to use the classical information, specifically to favour paths that are consistent with errors occurring where "bad" stabilizers took place, then the threshold improves (this is the case shown in the third panel of Fig. 4). In effect we trade a small amount of increased error risk for a significant amount of classical knowledge: the 4% of "bad" stabilizers account for about half of all errors entering the system.

MEMORY ERRORS

Since our protocols are considerably longer than the circuit required for the monolithic architecture, and indeed our approach is post-selective and therefore of uncertain duration, it is important to assess the potential impact of memory errors. In the main paper we assert that memory errors should have negligible impact, if our cells can employ qubits at least as good as those demonstrated in the *Science* papers of Maurer *et al* and Steger *et al*. To substantiate this we need to estimate the duration of the protocols; here we do so for EXPEDIENT since it is the protocol designed for use when memory errors are an issue. In the following we assign one "time step" to any elementary operation in our protocol, whether a gate, a raw Bell creation or a measurement. We assume that the operations within a cell must be strictly sequential. Because we are evaluating an entire 'sheet' of stabilizers in parallel across the array (see Fig. 3 of the main paper), we need to be concerned not merely with the average time that a stabilizer protocol might take, but rather with the the time required for a given target proportion of all stabilizers to succeed. Our simulations indicate that if we wait for 99% of stabilizers to evaluate, and abandon the remaining 1%, then there is negligible impact on the threshold. (A stabilizer that fails to report is less damaging than a stabilizer that performs a 'wrong' projection; moreover such stabilizers will not introduce errors since the ancilla GHZ state is never created and coupled to the data qubits.) Therefore we simply take

the expected time for 99% of stabilizer protocols to complete as the characteristic time for a 'sheet' of stabilizers to be found. Obviously there is scope for far more sophisticated approaches which minimise waiting time, but this naive strategy suffices to give us a bound.

The process of building a GHZ state can be separated into distinct sections, each of which is terminated by a measurement, which, if it results in the 'wrong' outcome, will reset the process to an earlier stage. Each of these measurement outcomes has an associated probability of success, p_L, where the index L denotes the level. For the EXPEDIENT protocol operating at error rates of $p_g = p_m = 0.6\%$ and $p_n = 10\%$, these levels and their probabilities are given in Table I. It should be noted that levels 1 and 2 utilise two cells (and 3,4 and 6,7) are run twice in parallel accross four cells, and consequently the longer of the two times will determine when the process may proceed to the next level.

The GHZ production process is probabalistic, with a minimum duration of $\sum_L t_L = 33$, which occurs with a probability $\prod_L p_L = 0.2242$. Using the parameters in the table the process was simulated; 100,000 samples were generated and used to estimate the parameters of the resulting distribution. This found the expectation time per stabilizer to be 68.2 time steps, while 50% of operations were completed after 57 time steps, 95% after 138 time steps, 99% after 195 time steps and 99.9% after 278 time steps.

Taking the value 195 steps, we find that 24 of these steps are manipulations of the memory qubits, i.e. a data qubit or a GHZ-bearing tier of the ancillas, which we are now taking to be nuclear spins. We can assume that conditional gates on these qubits are by far the slowest, at perhaps $10\mu s$ (see the Steger reference in the main paper). We therefore take $500\mu s$ as a characteristic time for a 99% subset of a complete 'sheet' of stabilizers to evaluate. Maurer *et al* reported that their best NV memory qubit had lifetimes of about 2 seconds (note the silicon impurity qubits were reported to live far longer, ~ 3 mins). If a proportion $1/e$ survive 2 seconds, we infer an error rate about 40% per second, or about 0.025% over $500\mu s$. This can indeed be neglected in comparison to the error rates from the active processes, i.e. the gates and measurements.